

# Introducción a la Teoría de Valores Extremos

## 3. Métodos Estadísticos

Joaquín Ortega Sánchez  
jortega@cimat.mx

Centro de Investigación en Matemáticas, CIMAT  
Guanajuato, Gto., Mexico

Octavo Congreso Latinoamericano de Sociedades de  
Estadística  
Montevideo, Octubre 2008

# Outline

Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

La Función de Excesos

Ejemplos

Uso de los Estadísticos de Orden

# Outline

## Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

La Función de Excesos

Ejemplos

Uso de los Estadísticos de Orden

# Introducción

Los métodos de estimación que estudiamos hasta ahora no hacen el mejor uso de la información disponible, pues de cada bloque sólo utilizamos un dato: el valor máximo.

Sin embargo, otros valores altos en el mismo bloque pueden tener información útil sobre la cola de la distribución de la muestra, que deberíamos tener en cuenta. Si tenemos información sobre la serie completa de datos, una alternativa es dejar de lado el procedimiento de bloques y estudiar todos los valores altos de la muestra.

# Introducción

El método de valores sobre un umbral fue propuesto inicialmente en el área de Hidrología y luego formalizado por R.L. Smith en base a resultados obtenidos previamente por Pickands.

La idea es seleccionar un umbral alto y considerar los valores de la muestra que están por encima de este umbral.

# Outline

Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

La Función de Excesos

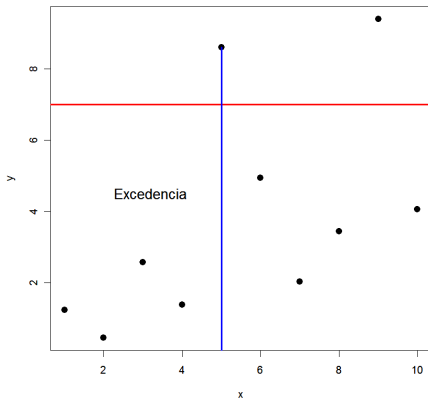
Ejemplos

Uso de los Estadísticos de Orden

# DGP

## Definición

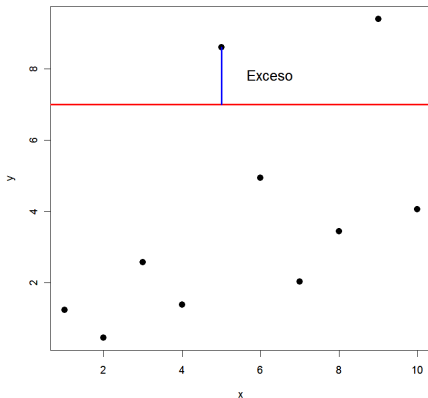
Sea  $X$  una v.a. con f.d.  $F$  y extremo derecho  $\omega_F$ . Para  $u < \omega_F$  fijo, decimos que ha ocurrido una excedencia de  $u$  si  $X > u$ . Llamamos excedencia al valor de  $X$  y exceso a  $X - u$ .



# DGP

## Definición

Sea  $X$  una v.a. con f.d.  $F$  y extremo derecho  $\omega_F$ . Para  $u < \omega_F$  fijo, decimos que ha ocurrido una excedencia de  $u$  si  $X > u$ . Llamamos excedencia al valor de  $X$  y exceso a  $X - u$ .





# DGP

## Definición

Llamaremos  $F_u$  a la fd de los excesos y  $F_{[u]}$  a la fd de las excedencias, que definimos como

$$\begin{aligned} F_u(x) &= F_{[u]}(x + u) = P(X - u \leq x | X > u) \\ &= \frac{F(x + u) - F(u)}{1 - F(u)}, \quad x \geq 0, \end{aligned} \quad (1)$$

La función

$$e(u) = E(X - u | X > u)$$

se conoce como la función media de excesos.

$F_u$  se conoce en otras áreas como la distribución de la *vida residual*, del *exceso de vida* o del *exceso de pérdida*.

# DGP

## Definición

Definimos la Distribución Generalizada de Pareto estándar (DGP)  $H_\xi$  por

$$H_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x} & \text{si } \xi = 0. \end{cases} \quad (2)$$

donde

$$\begin{aligned} x &\geq 0 && \text{si } \xi \geq 0, \\ 0 \leq x &\leq -1/\xi && \text{si } \xi < 0. \end{aligned}$$

Podemos también introducir una familia de ubicación y escala  $H_{\xi;\nu,\beta}$  reemplazando el argumento  $x$  por  $(x - \nu)/\beta$  para  $\nu \in \mathbb{R}$ ,  $\beta > 0$ . Hay que modificar en cada caso el soporte.

# DGP

Como para las *DVE*,  $H_0$  puede interpretarse como el límite de  $H_\xi$  cuando  $\xi \rightarrow 0$ . La f.d.  $H_{\xi;0,\beta}$  jugará un papel importante en el futuro, y para simplificar la notación escribiremos

$$H_{\xi,\beta}(x) = 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi}, \quad x \in D(\xi, \beta), \quad (3)$$

donde

$$D(\xi, \beta) = \begin{cases} [0, \infty) & \text{si } \xi \geq 0, \\ [0, -\beta/\xi] & \text{si } \xi < 0. \end{cases}$$

# DGP

Al igual que en el caso de las DGVE, hay otra parametrización posible para las DPG, separando en tres familias:

Tipo I: Exponencial  $W_0(x) = 1 - e^{-x}, \quad x \geq 0,$

Tipo II: Pareto  $W_{1,\alpha}(x) = 1 - x^{-\alpha}, \quad x \geq 1, \alpha > 0,$

Tipo III: Beta  $W_{2,\alpha}(x) = 1 - (-x)^\alpha, \quad -1 \leq x \leq 0, \alpha > 0.$

## Propiedades de la DGP

- (a) Supongamos que  $X \sim H_{\xi, \beta}$ . Entonces  $E(X) < \infty$  si y sólo si  $\xi < 1$ . En este último caso

$$E \left( 1 + \frac{\xi}{\beta} X \right)^{-r} = \frac{1}{1 + \xi r}, \quad r > -1/\xi,$$

$$E \left( \log \left( 1 + \frac{\xi}{\beta} X \right) \right)^k = \xi^k k!, \quad k \in \mathbb{N},$$

$$E \left( X(\overline{H}_{\xi, \beta}(X))^r \right) = \frac{\beta}{(r+1-\xi)(r+1)}, \quad \frac{r+1}{|\xi|} > 0.$$

Si  $\xi < 1/r$  con  $r \in \mathbb{N}$ , entonces

$$E[X^r] = \frac{\beta^r}{\xi^{r+1}} \frac{\Gamma(\xi^{-1} - r)}{\Gamma(1 + \xi^{-1})} r!.$$

## Propiedades de la DGP

(b) Para  $\xi \in \mathbb{R}$ ,  $F \in \mathcal{D}(G_\xi)$  si y sólo si

$$\lim_{u \uparrow \omega_F} \sup_{0 < x < \omega_F - u} |F_u(x) - H_{\xi, \beta(u)}(x)| = 0 \quad (4)$$

para alguna función positiva  $\beta$ .

Esta propiedad dice que las DGP son aproximaciones adecuadas para la f.d. de excesos  $F_u$  para  $u$  grande. Este resultado se debe a Pickands y puede reformularse como sigue. Para alguna función  $\beta$  que se estima a partir de los datos,

$$\bar{F}_u(x) = P(X - u > x | X > u) \approx \bar{H}_{\xi, \beta(u)}(x), \quad x > 0.$$

Es necesario que  $u$  sea suficientemente grande.

## Propiedades de la DGP

(c) Supongamos que  $x_i \in D(\xi, \beta)$ ,  $i = 1, 2$ , entonces

$$\frac{\overline{H}_{\xi, \beta}(x_1 + x_2)}{\overline{H}_{\xi, \beta}(x_1)} = \overline{H}_{\xi, \beta + \xi x_1}(x_2). \quad (5)$$

Esta propiedad se puede reformular diciendo que la clase de las DGP es cerrada respecto a cambios de umbral. El lado izquierdo de (5) es la probabilidad condicional de que, dado que la variable en consideración está por encima de  $x_1$ , también está por encima de  $x_1 + x_2$ . El lado derecho dice que esta probabilidad también es una DGP.

## Propiedades de la DGP

- (d) Sea  $N \sim \text{Pois}(\lambda)$ , independiente de la sucesión i.i.d.  $(X_n)$  con DGP de parámetros  $\xi$  y  $\beta$ . Escribimos  $M_N = \max(X_1, \dots, X_N)$ . Entonces

$$P(M_N \leq x) = \exp \left\{ -\lambda \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi} \right\} = G_{\xi; \mu, \psi}(x),$$

donde  $\mu = \beta \xi^{-1} (\lambda^\xi - 1)$  y  $\psi = \beta \lambda^\xi$ .

Esta propiedad dice que en un modelo en el cual el número de excedencias es exactamente Poisson y la f.d. de excesos es exactamente una DGP, el máximo de estos excesos tiene como distribución una DGVE.



## Propiedades de la DGP

- (e) Supongamos que  $X$  tiene una DGP con parámetros  $\xi < 1$  y  $\beta$ . Entonces para  $u < \omega_F$ ,

$$e(u) = E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad \beta + \xi u > 0.$$

En conjunto, (b) y (e) proveen una técnica gráfica para escoger una umbral  $u$  suficientemente alto como para que la aproximación de la f.d. de excesos  $F_u$  por una DGP se justifique:

# Propiedades de la DGP

Dada una muestra i.i.d.  $X_1, \dots, X_n$  construimos una función media de excesos empírica  $e_n(u)$  como una versión muestral de la función media de excesos  $e(u)$ . A partir de (e) vemos que la función media de excesos de una DGP es lineal, y entonces buscamos una región de valores de  $u$  donde la gráfica de  $e_n(u)$  sea aproximadamente lineal. Para estos valores de  $u$  parece razonable aproximar por una DGP.

# Outline

Introducción

La Distribución Generalizada de Pareto

**Excesos sobre un Umbral**

La Función de Excesos

Ejemplos

Uso de los Estadísticos de Orden

## Excesos sobre un Umbral

Supongamos que  $X_1, \dots, X_n$  son iid con fd  $F \in \mathcal{D}(G_\xi)$  para algún  $\xi \in \mathbb{R}$ . Primero escogemos un nivel alto  $u$  y llamamos

$$N_u = \text{card}\{i : i = 1, \dots, n : X_i > u\}$$

al número de excedencias de  $u$  por  $X_1, \dots, X_n$ . Los excesos correspondientes los denotamos por  $Y_1, \dots, Y_{N_u}$ . La fd de los excesos de  $X$  está dada por

$$F_u(y) = P(X - u \leq y | X > u) = P(Y \leq y | X > u), \quad y \geq 0.$$

Esta última relación también puede escribirse como

$$\bar{F}(u + y) = \bar{F}(u)\bar{F}_u(y). \quad (6)$$

## Excesos sobre un Umbral

Recordemos ahora la definición de la distribución generalizada de Pareto (DGP): Una DGP  $H_{\xi,\beta}$  con parámetros  $\xi \in \mathbb{R}$  y  $\beta > 0$  tiene cola

$$\bar{H}_{\xi,\beta}(x) = \begin{cases} \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi} & \text{si } \xi \neq 0, x \in D(\xi, \beta), \\ e^{-x/\beta} & \text{si } \xi = 0, x \in D(\xi, \beta), \end{cases}$$

donde

$$D(\xi, \beta) = \begin{cases} [0, \infty) & \text{si } \xi \geq 0, \\ [0, -\beta/\xi] & \text{si } \xi < 0. \end{cases}$$

## Excesos sobre un Umbral

Por la propiedad (b) sabemos que

$$\lim_{u \uparrow \omega_F} \sup_{0 < x < \omega_F - u} |\bar{F}_u(x) - \bar{H}_{\xi, \beta(u)}(x)| = 0,$$

para una función positiva adecuada  $\beta$ . Con base en este resultado, para  $u$  grande, podemos considerar la siguiente aproximación:

$$\bar{F}_u(x) \approx \bar{H}_{\xi, \beta(u)}(x) \quad (7)$$

Es importante observar que  $\beta$  es una función del umbral  $u$ . Si  $F \in \mathcal{D}(G_{\xi, \mu, \sigma})$  entonces  $\beta = \sigma + \xi(u - \mu)$ . En la práctica,  $u$  debe ser suficientemente grande. Dado este  $u$  podemos estimar  $\xi$  y  $\beta = \beta(u)$  a partir de los datos de los excesos, de modo que las estimaciones que resultan dependen de  $u$ .

## Excesos sobre un Umbral

A continuación damos las ideas principales de la justificación de (7). Supongamos que  $X$  tiene distribución  $F \in \mathcal{D}(G_{\xi, \mu, \sigma})$ . Para  $n$  suficientemente grande

$$F^n(z) \approx \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}.$$

En consecuencia

$$n \log F(z) \approx - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}.$$

Pero para valores grandes de  $z$ , haciendo un desarrollo de Taylor se obtiene que

$$\log F(z) \approx -(1 - F(z))$$

y usando esta relación en la ecuación anterior obtenemos

$$1 - F(u) \approx \frac{1}{n} \left[ 1 + \xi \left( \frac{u - \mu}{\sigma} \right) \right]^{-1/\xi}.$$

para  $u$  grande.

## Excesos sobre un Umbral

De manera similar, para  $y > 0$ ,

$$1 - F(u + y) \approx \frac{1}{n} \left[ 1 + \xi \left( \frac{u + y - \mu}{\sigma} \right) \right]^{-1/\xi}.$$

En consecuencia

$$\begin{aligned} P(X > u + y | X > u) &\approx \frac{n^{-1} [1 + \xi(u + y - \mu)/\sigma]^{-1/\xi}}{n^{-1} [1 + \xi(u - \mu)/\sigma]^{-1/\xi}} \\ &= \left[ \frac{1 + \xi(u + y - \mu)/\sigma}{1 + \xi(u - \mu)/\sigma} \right]^{-1/\xi} \\ &= \left[ 1 + \xi \frac{y}{\beta} \right]^{-1/\xi}, \end{aligned}$$

donde  $\beta = \sigma + \xi(u - \mu)$ .



## Excesos sobre un Umbral

### Ejemplos: Distribución Exponencial

Para el modelo exponencial  $F(x) = 1 - e^{-x}$  para  $x > 0$ . Por lo tanto

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{e^{-(u+y)}}{e^{-u}} = e^{-y}$$

para todo  $y > 0$ . En consecuencia la distribución límite para las excedencias es la distribución exponencial, que corresponde a  $\xi = 0$  y  $\sigma' = u$ . Este es un resultado exacto para todos los umbrales  $u > 0$

## Excesos sobre un Umbral

### Ejemplos: Distribución de Fréchet

La distribución de Fréchet estándar es  $F(x) = \exp(-1/x)$ , para  $x > 0$ . Por lo tanto

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{1 - e^{\{-(u+y)^{-1}\}}}{1 - e^{-u^{-1}}} \sim \left(1 + \frac{y}{u}\right)^{-1}$$

cuando  $u \rightarrow \infty$ , para todo  $y > 0$ . Esto corresponde a una distribución generalizada de Pareto con  $\xi = 1$  y  $\sigma' = u$

## Excesos sobre un Umbral

### Ejemplos: Distribución Uniforme

La función de distribución uniforme en  $[0, 1]$  es  $F(x) = x$ , para  $0 \leq x \leq 1$ . Por lo tanto

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{1 - (u + y)}{1 - u} = 1 - \frac{y}{1 - u}$$

para  $0 \leq y \leq 1 - u$ . Esto corresponde a una distribución generalizada de Pareto con  $\xi = -1$  y  $\sigma' = 1 - u$

## Excesos sobre un Umbral

La relación (6) sugiere un método para estimar la parte más lejana de la cola de  $F$ , estimando  $\bar{F}_u(y)$  y  $\bar{F}(u)$  por separado. Un estimador natural de  $\bar{F}(u)$  está dado por la fde

$$\hat{\bar{F}}(u) = \bar{F}_n(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i > u\}} = \frac{N_u}{n}.$$

Por otro lado, la aproximación por la DGP (7) sugiere un estimador de la forma

$$\hat{\bar{F}}_u(y) = \bar{H}_{\hat{\xi}, \hat{\beta}}(y) \quad (8)$$

para estimadores apropiados  $\hat{\xi} = \hat{\xi}_{N_u}$  y  $\hat{\beta} = \hat{\beta}_{N_u}$ .

## Excesos sobre un Umbral

El estimador resultante para la cola  $\bar{F}(u + y)$  para  $y > 0$  tiene la forma

$$\hat{\bar{F}}(u + y) = \frac{N_u}{n} \left( 1 + \hat{\xi} \frac{y}{\hat{\beta}} \right)^{-1/\hat{\xi}}. \quad (9)$$

En el caso de las distribuciones de Fréchet y Gumbel ( $\xi \geq 0$ ), el dominio en (9) es  $y \geq 0$  y obtenemos como estimador del cuantil

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left( \left( \frac{n}{N_u} (1 - p) \right)^{-\hat{\xi}} - 1 \right). \quad (10)$$

Además, para  $\hat{\xi} < 0$  un estimador del extremo derecho  $\hat{\omega}_F$  de  $F$  es

$$\hat{\omega}_F = u - \frac{\hat{\beta}}{\hat{\xi}}.$$

Este valor se obtiene poniendo  $\hat{\omega}_F = \hat{x}_1$  en (10).

## Excesos sobre un Umbral

Si deseamos escoger un umbral óptimo  $u$  nos encontramos con problemas usuales de sesgo y varianza. Un valor demasiado alto de  $u$  resulta en pocas excedencias y en consecuencia estimadores de varianza alta. Para  $u$  demasiado pequeño los estimadores son sesgados.

Un método de selección disponible en la práctica está basado en la linealidad de la función media de excesos  $e(u)$  para la DGP, como mencionamos anteriormente.

## Excesos sobre un Umbral

### Estimación por Máxima Verosimilitud.

Los siguientes resultados se deben a Smith. Recordemos que los datos originales  $\mathbf{X} = (X_1, \dots, X_n)$  son iid con fd  $F$ .

Supongamos que  $F$  es DGP con parámetros  $\xi$  y  $\beta$ , de modo que la densidad  $f$  es

$$f(x) = \frac{\xi}{\beta} \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}-1}, \quad x \in D(\xi, \beta).$$

La función de log-verosimilitud es

$$\ell((\xi, \beta); \mathbf{X}) = -n \log \beta - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\beta} X_i\right).$$

Los argumentos de la función de verosimilitud deben satisfacer que  $X_i \in D(\xi, \beta)$ . Recordemos que  $D(\xi, \beta) = [0, \infty)$  para  $\xi \geq 0$ .

## Excesos sobre un Umbral

Las ecuaciones de verosimilitud pueden ser derivadas y resueltas numéricamente, de donde se obtienen los EMV  $\hat{\xi}_n, \hat{\beta}_n$ . Este método funciona bien si  $\xi > -1/2$ , y en este caso se puede mostrar que

$$n^{1/2} \left( \hat{\xi}_n - \xi, \frac{\hat{\beta}_n}{\beta} - 1 \right) \xrightarrow{d} \mathcal{N}(0, M^{-1}), \quad n \rightarrow \infty,$$

donde

$$M^{-1} = (1 + \xi) \begin{pmatrix} 1 + \xi & 1 \\ 1 & 2 \end{pmatrix}$$

También valen las propiedades usuales de consistencia y eficiencia asintótica.



## Excesos sobre un Umbral

### Momentos Pesados por Probabilidad.

Hoskins & Wallis obtuvieron estimadores de este tipo para la DGP, que están basados en

$$w_r = E[Z(\bar{H}_{\xi, \beta}(Z))^r] = \frac{\beta}{(r+1)(r+1-\xi)}, \quad r = 0, 1,$$

donde  $Z$  tiene DGP  $H_{\xi, \beta}$ . De aquí obtenemos

$$\beta = \frac{2w_0w_1}{w_0 - 2w_1} \quad \text{y} \quad \xi = 2 - \frac{w_0}{w_0 - 2w_1}.$$

Si reemplazamos ahora  $w_0$  y  $w_1$  por sus estimadores empíricos, obtenemos los estimadores de momentos pesados por probabilidades  $\hat{\beta}$  y  $\hat{\xi}$ .

## Excesos sobre un Umbral

### La Gráfica de Cuantiles (qq-plot) Pareto

Esperamos que el qq-plot Pareto será eventualmente lineal si la cola de la distribución  $1 - F$  es de tipo Pareto y además, la pendiente de esta parte lineal del qq-plot Pareto es el índice  $\xi = 1/\alpha$ . Para estimar esta pendiente podemos usar el valor muestral  $J_{k,n}$  de la función media de excesos del logaritmo de las variables aleatorias, que se define por

$$J_{k,n} = \frac{1}{k-1} \sum_{j=1}^{k-1} \log X_{[j,n]} - \log X_{[k,n]}$$

$J_{k,n}$  puede interpretarse como el aumento promedio del qq-plot Pareto a partir de cierto punto  $(-\log(\frac{k+1}{n+1}), \log X_{[k+1,n]})$  en adelante.

## Excesos sobre un Umbral

Otra alternativa es considerar un enfoque de regresión para el problema de estimación: queremos hallar una recta de regresión en el qq plot que pase por el punto  $(-\log(\frac{k+1}{n+1}), \log X_{[k+1,n]})$  y que trate de ajustarse a la gráfica de los puntos

$$\left(-\log\left(\frac{j}{n+1}\right), \log X_{[j,n]}\right), \quad j = 1, \dots, k.$$

Si hacemos una regresión con un algoritmo de mínimos cuadrados pesados tenemos que minimizar

$$\sum_{j=1}^k w_{j,k} \left( \log X_{[j,n]} - \left( \log X_{[k+1:n]} + g \log\left(\frac{k+1}{j}\right) \right) \right)^2.$$

respecto a  $g$  ( $g > 0$ ) para ciertos pesos  $w_{j,k}$ ,  $1 \leq j \leq k$  y  $k \geq 1$ .

## Excesos sobre un Umbral

Derivando respecto a  $g$  se obtiene el siguiente estimador de  $\xi$

$$\hat{\xi} = \frac{\sum_{j=1}^k w_{j,k} \log((k+1)/j) (\log X_{[j,n]} - \log X_{[k+1,n]})}{\sum_{j=1}^k w_{j,k} \log^2(j/(k+1))}$$

Poniendo  $\tilde{w}_{j,k} = w_{j,k} \log(\frac{k+1}{j})$  obtenemos

$$\begin{aligned} \frac{\hat{\xi}}{k} \sum_{j=1}^k w_{j,k} \log^2\left(\frac{j}{k+1}\right) &= \frac{1}{k} \sum_{j=1}^k \tilde{w}_{j,k} (\log X_{[j,n]} - \log X_{[k+1,n]}) \\ &= \sum_{m=1}^k \frac{m}{k} \left( \frac{1}{m} \sum_{j=1}^m \tilde{w}_{j,k} \right) (\log X_{[m,n]} - \log X_{[m+1,n]}) \end{aligned}$$

## Excesos sobre un Umbral

Si escribimos  $w_{j,k}$  como  $w(j/k)$  e introducimos

$K(t) = -\frac{1}{t} \int_0^t w(x) \log x \, dx$ , entonces  $\frac{1}{m} \sum_{j=1}^k \tilde{w}_{j,k}$  es una aproximación de  $K(m/k)$ . Es fácil ver que el estimador  $J_{k,n}$  corresponde al caso en el cual  $w_{j,k} = 1/\log(\frac{k+1}{j})$ ,  $j = 1, \dots, k$  ó  $\tilde{w}_{j,k} = 1$ .

## Excesos sobre un Umbral

Concluimos que el método de estimación de mínimos cuadrados pesados para la pendiente del qq-plot lleva a un estimador tipo núcleo (kernel)

$$\hat{\xi}_K = \frac{\sum_{m=1}^k \frac{m}{k} K\left(\frac{m}{k}\right) (\log X_{[m,n]} - \log X_{[m+1,n]})}{k^{-1} \sum_{m=1}^k K\left(\frac{m}{k}\right)} \quad (11)$$

Se puede ver que el denominador de (11) es una aproximación de la integral  $\int_0^1 K(t) dt$ .

Estos estimadores fueron propuestos inicialmente por Csörgő, Deheuvels y Mason (1985), quienes probaron su consistencia, y además probaron la normalidad asintótica de los estimadores de núcleo bajo condiciones generales.

# Pruebas de Hipótesis

## Prueba basada en el coeficiente de variación muestral.

El coeficiente de variación es la desviación típica dividida por la media. El recíproco del coeficiente de variación es igual a  $1 - 2\xi$  para DGP  $H_{\xi,\beta}$  con  $\xi < 1/2$ . Si  $y_1, \dots, y_k$  son las excedencias sobre el umbral  $u$  podemos usar el siguiente estadístico para la prueba  $(\bar{y}_k - u)^2 / s_k^2$ , que es invariante bajo cambios de escala. Usando los valores ordenados obtenemos el siguiente estadístico para la prueba

$$\frac{(\bar{x}_k - x_{[k,n]})^2}{\frac{1}{k-1} \sum_{i=1}^k (x_{[i,k]} - \bar{x}_k)^2}$$

donde  $\bar{x}_k = \frac{1}{k} \sum_{i \leq k} x_{[i,n]}$ . Este estadístico fue introducido por Hashofer & Wang para probar  $\xi = 0$  contra  $\xi \neq 0$ .

# Pruebas de Hipótesis

## Prueba del cociente de verosimilitudes para el modelo exponencial

Para hacer una prueba de

$$H_0 : \xi = 0 \quad \text{contra} \quad H_1 : \xi \neq 0$$

con parámetros de forma desconocidos  $\beta > 0$ , dado un vector  $\mathbf{y} = (y_1, \dots, y_k)$  de excedencias sobre el umbral  $u$ , el estadístico del cociente de verosimilitudes es

$$T_{LR}(\mathbf{y}) = 2 \log \frac{\prod_{i \leq k} h_{\hat{\xi}, u, \hat{\beta}}(y_i)}{\prod_{i \leq k} h_{0, u, \tilde{\beta}}(y_i)}$$

donde  $(\hat{\xi}, \hat{\beta})$  y  $\tilde{\beta}$  denotan los EMV para los modelos DGP y DGP0. El  $p$ -valor es

$$p_{LR}(\mathbf{y}) = 1 - \chi_1^2(T_{LR}(\mathbf{y})).$$



# Outline

Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

**La Función de Excesos**

Ejemplos

Uso de los Estadísticos de Orden

## La Función de Excesos

Recordamos que definimos la función media de excesos como

$$e(u) = E(X - u | X > u), \quad 0 \leq u < \omega_F, \quad (12)$$

que se puede calcular usando la fórmula

$$e(u) = \frac{1}{\bar{F}(u)} \int_u^{\omega_F} \bar{F}(x) dx, \quad 0 < u < \omega_F. \quad (13)$$

y en el teorema 2.18 vimos que si  $X$  tiene una DGP con parámetros  $\xi < 1$  y  $\beta$  entonces

$$e(u) = \frac{\beta + \xi u}{1 - \xi}, \quad \beta + \xi u > 0.$$

## La Función de Excesos

### Ejemplo

Si  $X \sim \text{Exp}(\lambda)$  entonces  $e(u) = \lambda^{-1}$  para todo  $u > 0$ :

$$e(u) = \frac{1}{e^{-\lambda u}} \int_u^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

### Ejemplo

Supongamos que  $X$  es una v.a. con fd  $F$  y  $\omega_F = \infty$ . Si para todo  $y \in \mathbb{R}$ ,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x-y)}{\bar{F}(x)} = e^{\gamma y}, \quad (14)$$

para algún  $\gamma \in [0, \infty]$ , entonces

$$\lim_{u \rightarrow \infty} e(u) = \gamma^{-1}$$

# La Función de Excesos

## Ejemplo

*Recordemos que si  $X$  tiene DGP la función media de excesos es lineal. La función media de excesos de una fd de colas pesadas, para valores grandes del argumento, está típicamente entre una constante (para  $\text{Exp}(\lambda)$ ) y una recta con pendiente positiva (para el caso Pareto). En consecuencia, las funciones medias de excesos interesantes son de la forma*

$$e(u) = \begin{cases} u^{1-\beta}/\alpha, & \alpha > 0, 0 \leq \beta < 1, \\ u/(\alpha + 2\beta \log u), & \alpha, \beta > 0. \end{cases}$$

*Observamos que  $e(u)$  aumenta con  $u$  pero la tasa de incremento disminuye con  $u$ .*

## La Función de Excesos

Es posible basar un procedimiento gráfico para estudiar el comportamiento de las colas en la función media de excesos empírica  $e_n(u)$ . Supongamos que  $X_1, \dots, X_n$  son i.i.d. con fd  $F$  y sea  $F_n$  la fde y  $\Delta_n(u) = \{i : i = 1, \dots, n, X_i > u\}$ , entonces

$$e_n(u) = \frac{1}{\overline{F}_n(u)} \int_u^\infty (x - u) dF_n(y) \quad (15)$$

$$= \frac{1}{\text{card}\Delta_n(u)} \sum_{i \in \Delta_n(u)} (X_i - u), \quad u \geq 0, \quad (16)$$

con la convención  $0/0 = 0$ . Una gráfica de la función media de excesos es una gráfica de los puntos

$$(X_{k:n}, e_n(X_{k:n})), \quad k = 1, \dots, n.$$

## La Función de Excesos

Observamos que

$$e_n(X_{k:n}) = \frac{1}{n-k} \sum_{j=1}^{n-k} X_{n-j+1:n} - X_{k:n}$$

Usamos la gráfica de la función media de excesos principalmente para distinguir entre modelos de colas livianas y pesadas. Es necesario tener precaución a la hora de interpretar estas gráficas. Debido a que hay pocos datos disponibles para calcular  $e_n(u)$  para valores grandes de  $u$ , las gráficas son muy sensibles a cambios en los datos hacia el final del rango.

## La Función de Excesos

Para tratar de hacer este procedimiento más robusto, se han propuesto las siguientes variaciones de  $e_n(X_{k:n})$

- La función media de excesos recortada  $T_{k,n}^{(p)}$  para una cierta proporción de recorte  $p \in (0, 1)$ ,

$$T_{k,n}^{(p)} = \frac{1}{k - \lfloor pk \rfloor} \sum_{\lfloor pk \rfloor + 1}^k X_{n-j+1:n} - X_{n-k:n}$$

donde  $\lfloor x \rfloor$  denota el mayor entero que no es mayor que  $x$ .

- Otra propuesta es reemplazar el promedio de los datos que son mayores que  $X_{k:n}$  por una mediana generalizada

## La Función de Excesos

En cuanto a la selección del umbral  $u$ , sabemos que para una va con fd  $H_{\xi, \beta}$ ,

$$e(u) = E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad u \in D(\xi, \beta), \xi < 1,$$

y por lo tanto  $e(u)$  es lineal. Recordemos que la función media de excesos empírica de la muestra  $X_1, \dots, X_n$  es

$$e_n(u) = \frac{1}{N_u} \sum_{i \in \Delta_n(u)} (X_i - u), \quad u > 0,$$

donde  $N_u = \text{card}\{i : i = 1, \dots, n, X_i > u\} = \text{card}\Delta_n(u)$ .

Entonces, podemos escoger  $u > 0$  de modo que  $e_n(u)$  sea aproximadamente lineal para  $x \geq u$ .



# Outline

Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

La Función de Excesos

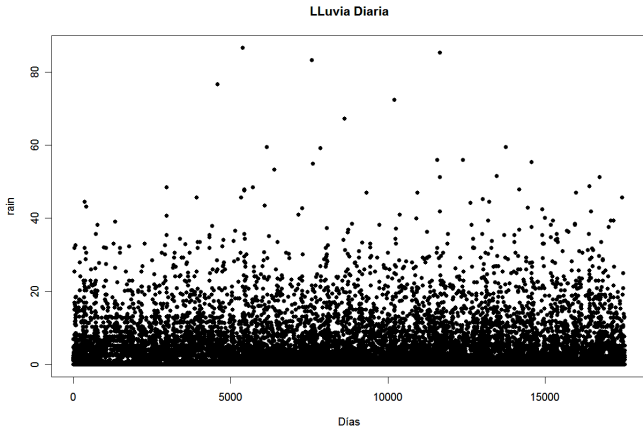
**Ejemplos**

Uso de los Estadísticos de Orden

# Ejemplos: Lluvias

## Registros diarios de lluvias

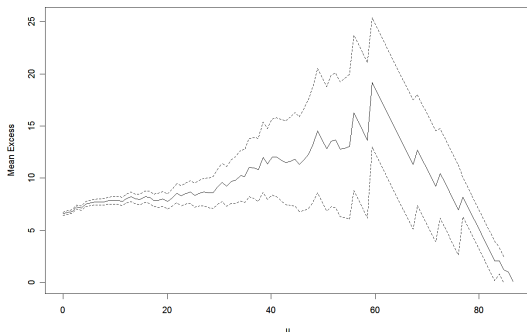
La figura muestra una serie de tiempo de las lluvias diarias en un estación de Inglaterra para el periodo 1914-1962. Hay un total de 17,531 datos.



## Ejemplos: Lluvias

Si optamos por el método de máximos por año usaríamos únicamente 48 datos, el 0.0028% de los datos.

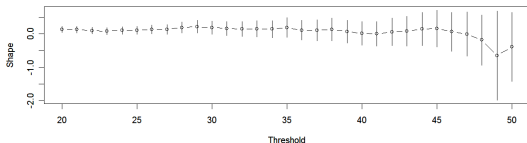
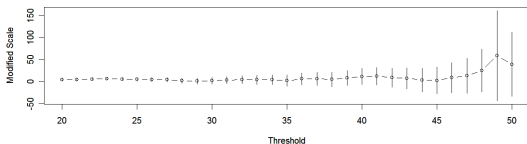
Vamos a estudiar el problema usando el enfoque de excesos de un umbral y para seleccionar el umbral hacemos una gráfica de la función de excesos.



## Ejemplos: Lluvias

El gráfico anterior no tiene una interpretación clara pero sugiere iniciar con un umbral alrededor de 30.

Hacemos ahora una gráfica de los valores estimados de  $\xi$  y  $\beta$  como función del umbral para obtener más información.



## Ejemplos: Lluvias

Esta gráfica confirma que un umbral de 30 es una selección razonable. Usando este valor obtenemos los EMV

$$(\hat{\xi}, \hat{\beta}) = (0.184, 7.44)$$

con un valor para la log-verosimilitud de -485.1 La matriz de varianzas y covarianzas es

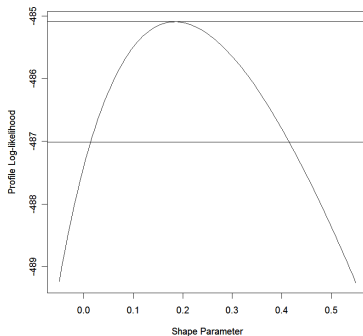
$$\begin{pmatrix} 0.0102 & -0.0655 \\ -0.0655 & 0.9188 \end{pmatrix}$$

con errores estándar de 0.101 y 0.959 para  $\hat{\xi}$  y  $\hat{\beta}$ , respectivamente.

Usando la normalidad asintótica de los estimadores obtenemos un intervalo de confianza de 95% para  $\xi$  de  $[-0.014, 0.383]$

## Ejemplos: Lluvias

Verosimilitud perfil para el parámetro de forma.



Intervalo de confianza asociado  $[0.019, 0.418]$

## Ejemplos: Lluvias

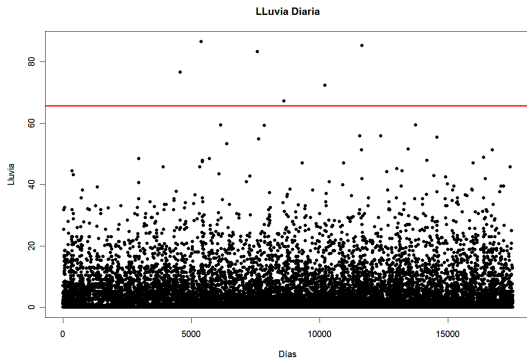
Hay 152 excedencias del nivel  $u = 30$  lo que da una estimación para la probabilidad de excederlo  $\hat{\zeta}_u = 152/17531 = 0.00867$  con varianza aproximada

$$\text{Var}(\hat{\zeta}_u) = \hat{\zeta}_u(1 - \hat{\zeta}_u)/17531 = 4.9 \times 10^{-7}$$

En cuanto a niveles de retorno, el nivel estimado correspondiente a 100 años es de 106.3 mm. con varianza 431.3, que lleva a un intervalo de [65.6, 147.0].

Una comparación con los datos muestra que este intervalo es dudoso. En los 48 años de observación el extremo inferior fue superado 6 veces, lo que sugiere que el nivel de retorno de 100 años debe estar por encima de estos valores.

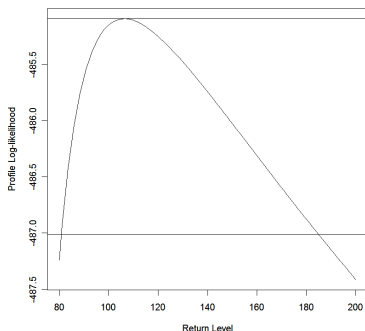
# Ejemplos: Lluvias





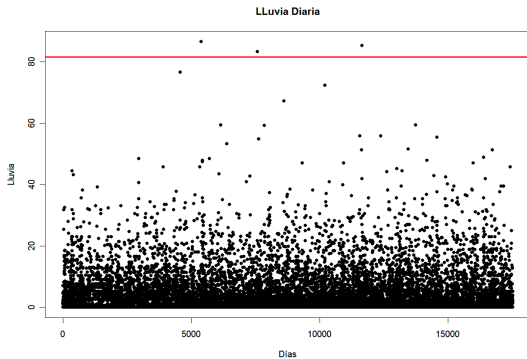
## Ejemplos: Lluvias

La figura muestra la verosimilitud perfil para el nivel de retorno de 100 años.

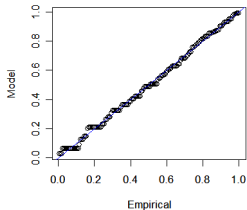
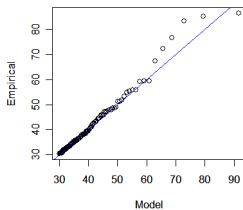
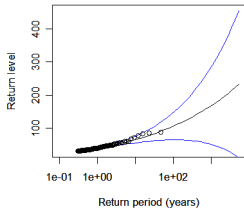
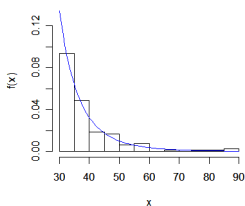


Intervalo de confianza aproximado al 95% es [81.6, 185.7]

# Ejemplos: Lluvias



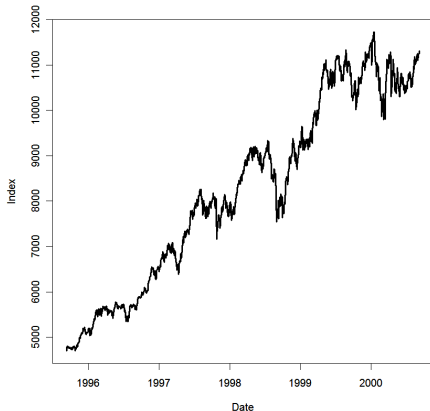
# Ejemplos: Lluvias

**Probability Plot****Quantile Plot****Return Level Plot****Density Plot**

# Ejemplos: Dow Jones

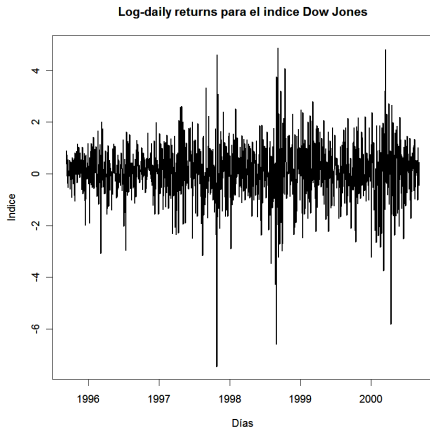
## Indice Dow Jones

La grafica muestra la evolución del índice Dow Jones durante un período de 5 años. Obviamente se trata de un proceso no estacionario.



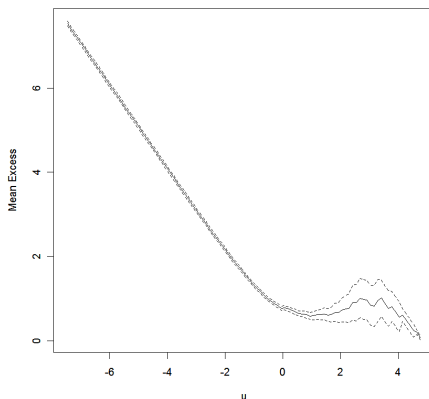
## Ejemplos: Dow Jones

Para obtener una aproximación a la estacionaridad tomamos logaritmos de los cocientes de observaciones sucesivas (daily log-returns). Los datos ahora son  $\log X_i - \log X_{i-1}$  y para mejorar la presentación multiplicamos estos valores por 100.



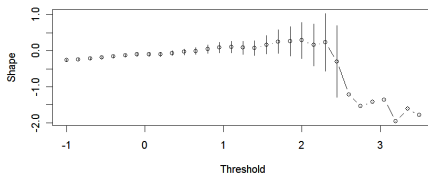
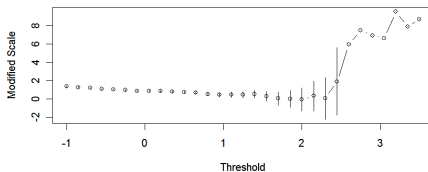
## Ejemplos: Dow Jones

Gráfica de la función media de excesos. Umbral sugerido:  
 $u = 2$ .



# Ejemplos: Dow Jones

Gráfica de la función media de excesos. Umbral sugerido:  
 $u = 1.5$ .



## Ejemplos: Dow Jones

Con  $u = 2$  tenemos 37 excedencias del nivel  $u$  en una serie de 1303 datos, de modo que la probabilidad de excedencia se estima en  $37/1303 = 0.028$ .

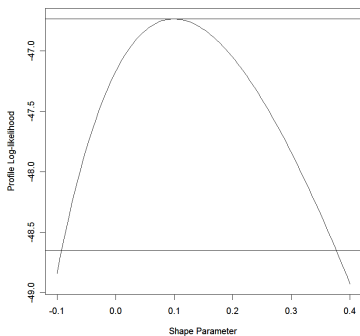
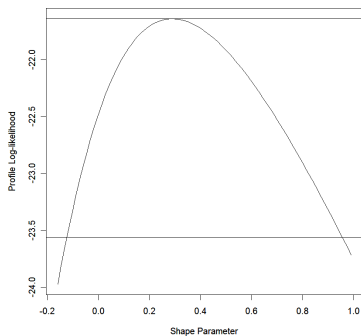
Los EMV para la DGP son  $(\hat{\xi}, \hat{\beta}) = (0.288, 0.495)$  con errores estándar  $(0.258, 0.150)$ . Para  $\xi$  tenemos un intervalo de confianza  $(-0.218, 0.794)$ , que incluye al 0.

Con  $u = 1.5$  tenemos 86 excedencias del nivel  $u$  en una serie de 1303 datos, de modo que la probabilidad de excedencia se estima en  $86/1303 = 0.066$ .

Los EMV para la DGP son  $(\hat{\xi}, \hat{\beta}) = (0.099, 0.573)$  con errores estándar  $(0.116, 0.09)$ . Para  $\xi$  tenemos un intervalo de confianza  $(-0.128, 0.327)$ , que también incluye al 0.

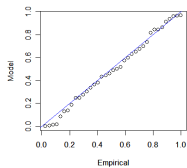


# Ejemplos: Dow Jones

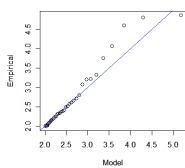


# Ejemplos: Dow Jones

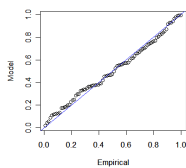
Probability Plot



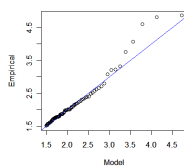
Quantile Plot



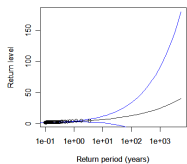
Probability Plot



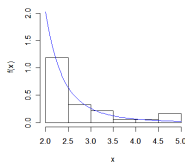
Quantile Plot



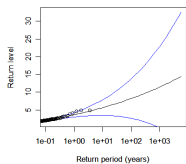
Return Level Plot



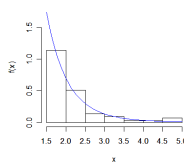
Density Plot



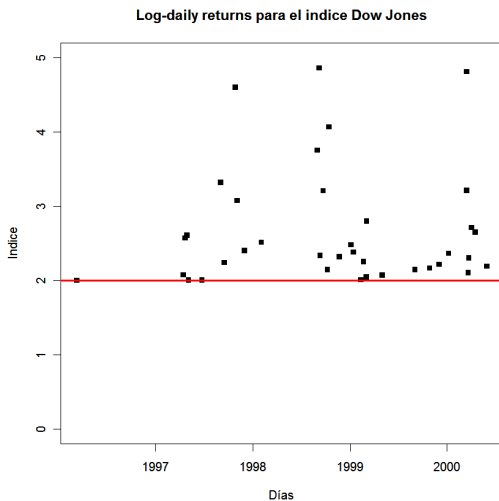
Return Level Plot



Density Plot



# Ejemplos: Dow Jones



# Outline

Introducción

La Distribución Generalizada de Pareto

Excesos sobre un Umbral

La Función de Excesos

Ejemplos

Uso de los Estadísticos de Orden

# Estadísticos de Orden

Una dificultad inherente a cualquier análisis de valores extremos es la cantidad limitada de datos disponibles para la estimación del modelo. Por definición los extremos son escasos, de modo que las estimaciones hechas a través de los modelos que se obtienen, especialmente para cuantiles altos de la distribución, tienen varianza grande. Ya hemos visto el método de excesos sobre un umbral, que permite ampliar la información a partir de la cual hacemos inferencia, de modo de tratar de mejorar la precisión de la estimación. Otra manera de hacer esto es estudiar el comportamiento de los estadísticos de orden superior de orden mayor o igual a  $r$  en un bloque dado, para valores pequeños de  $r$ .

# Estadísticos de Orden

Supongamos que  $X_1, X_2, \dots$  es una sucesión i.i.d. y queremos caracterizar su comportamiento extremal. Vimos que la distribución límite cuando  $n \rightarrow \infty$  de  $M_n$  adecuadamente normalizado es la DGVE. Inicialmente extendemos este resultado a otros estadísticos de orden. Recordamos que

$$X_{[k,n]} = k\text{-ésimo mayor de } \{X_1, \dots, X_n\}$$

y buscamos estudiar el comportamiento límite de esta variable, para  $k$  fijo, cuando  $n \rightarrow \infty$ . El siguiente resultado es una generalización del teorema de Fisher, Tippet y Gnedenko.

# Estadísticos de Orden

## Teorema

Si existen sucesiones  $a_n > 0$  y  $b_n$  tales que

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z), \quad n \rightarrow \infty,$$

para alguna f. d. no degenerada  $G$  entonces, para  $k$  fijo,

$$P\left(\frac{X_{[k,n]} - b_n}{a_n} \leq z\right) \rightarrow G_k(z), \quad n \rightarrow \infty,$$

en el conjunto  $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$ , donde

$$G_k(z) = \exp\{-\tau(z)\} \sum_{s=0}^{k-1} \frac{\tau^s(z)}{s!} \quad (17)$$

con  $\tau(z) = \left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}$

# Estadísticos de Orden

Este teorema implica que si el  $k$ -ésimo mayor estadístico de orden se normaliza exactamente igual que el máximo, su distribución límite tiene la forma descrita por la ecuación (17), cuyos parámetros corresponden a los parámetros de la DGVE límite para los máximos por bloque.

Sin embargo, el teorema sólo nos habla de las distribuciones unidimensionales de los estadísticos de orden, y no de su distribución conjunta.



## Estadísticos de Orden

En general, vamos a tener como datos un vector con los  $r$  mayores estadísticos de orden dentro de cada bloque:

$$\mathbf{M}_n^{(r)} = (X_{[1,n]}, \dots, X_{[r,n]}).$$

Está claro que las componentes de este vector no pueden ser independientes:  $X_{[2,n]}$  no puede ser mayor que  $X_{[1,n]} = M_n$ , por ejemplo. Por lo tanto el resultado del teorema anterior no nos da una distribución para el vector  $\mathbf{M}_n^{(r)}$ .

Con algo de trabajo y re-escalando apropiadamente, es posible obtener la distribución conjunta, pero la expresión que resulta es demasiado complicada para ser de utilidad. Sin embargo, el siguiente teorema nos da la densidad conjunta de la distribución límite:

# Estadísticos de Orden

## Teorema

*Si existen sucesiones de constantes  $a_n > 0$  y  $b_n$  tales que*

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z), \quad n \rightarrow \infty,$$

*para alguna función de distribución no degenerada  $G$ , entonces, para  $r$  fijo, la distribución límite cuando  $n \rightarrow \infty$  de*

$$\tilde{\mathbf{M}}_n^{(r)} = \left( \frac{X_{[1,n]} - b_n}{a_n}, \dots, \frac{X_{[r,n]} - b_n}{a_n} \right).$$

*está dentro de la familia de distribuciones con densidad conjunta*

## Estadísticos de Orden

$$f(z^{(1)}, \dots, z^{(r)}) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z^{(r)} - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (18)$$

$$\times \prod_{k=1}^r \sigma^{-1} \left[ - \left( \frac{z^{(k)} - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi} - 1}, \quad (19)$$

donde  $-\infty < \mu < \infty$ ,  $\sigma > 0$  y  $-\infty < \xi < \infty$ ;

$z^{(r)} \leq z^{(r-1)} \leq \dots \leq z^{(1)}$  y  $1 + \xi(z^{(k)} - \mu)/\sigma > 0$  para  $k = 1, \dots, r$ .

# Estadísticos de Orden

El caso  $\xi = 0$  se interpreta como el límite cuando  $\xi \rightarrow 0$  de (19):

$$f(z^{(1)}, \dots, z^{(r)}) = \exp \left\{ - \exp \left[ - \left( \frac{z^{(r)} - \mu}{\sigma} \right) \right] \right\} \\ \times \prod_{k=1}^r \sigma^{-1} \exp \left[ - \left( \frac{z^{(k)} - \mu}{\sigma} \right) \right], \quad (20)$$

# Estadísticos de Orden

Para modelar los estadísticos de orden comenzamos con una colección de v.a.i.i.d. agrupadas en  $m$  bloques. En el  $i$ -ésimo bloque se registran las mayores  $r_i$  observaciones y obtenemos la serie de vectores

$$\mathbf{M}_i^{(r_i)} = (z_i^{(1)}, \dots, z_i^{(r_i)}), \quad i = 1, \dots, m$$

Lo usual es poner todos los  $r_i$  iguales a un mismo valor  $r$ , a menos que algún bloque no tenga suficientes datos.

## Estadísticos de Orden

Al igual que para el método de máximos por bloques, la selección del tamaño de los bloques representa un compromiso entre sesgo y varianza, que se resuelve usualmente haciendo una selección que tenga sentido pragmático, como por ejemplo tomando bloques de tamaño un año. El número de estadísticos de orden que se usa en cada bloque también representa un compromiso entre sesgo y varianza, similar al de los excesos sobre un umbral. En la práctica seleccionamos los  $r_i$  tan grandes como sea posible, sin violar las hipótesis del modelo, según lo muestren las gráficas correspondientes.

## Estadísticos de Orden

La verosimilitud para este modelo se obtiene a partir de (19) y (20): para  $\xi \neq 0$ ,

$$L(\mu, \sigma, \xi) = \prod_{i=1}^m \left( \exp \left\{ - \left[ 1 + \xi \left( \frac{z_i^{(r_i)} - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \right) \quad (21)$$

$$\times \prod_{k=1}^{r_i} \sigma^{-1} \left[ 1 + \xi \left( \frac{z_i^{(k)} - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi} - 1}, \quad (22)$$

siempre que  $1 + \xi(z^{(k)} - \mu)/\sigma > 0$ ,  $k = 1, \dots, r_i$ ,  $i = 1, \dots, m$ .

## Estadísticos de Orden

Cuando  $\xi = 0$ ,

$$L(\mu, \sigma, \xi) = \prod_{i=1}^m \left( \exp \left\{ - \exp \left[ - \left( \frac{z_i^{(r_i)} - \mu}{\sigma} \right) \right] \right\} \right) \quad (23)$$

$$\times \prod_{k=1}^{r_i} \sigma^{-1} \left[ 1 + \xi \left( \frac{z_i^{(k)} - \mu}{\sigma} \right) \right], \quad (24)$$

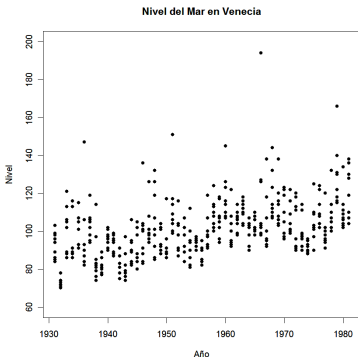
Las log-verosimilitudes correspondientes pueden ser maximizadas numéricamente para obtener EMV. La inclusión de una mayor cantidad de información debería mejorar la precisión de los estimadores.



# Estadísticos de Orden

## Ejemplo: Nivel del mar en Venecia.

Los datos corresponden a los 10 mayores estadísticos de orden para el nivel del mar en Venecia, salvo para el año 1935, para el cual sólo hay 6 observaciones disponibles.

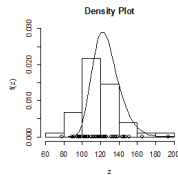
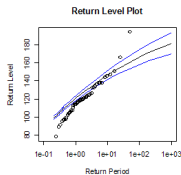
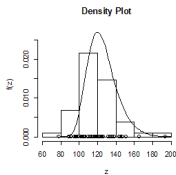
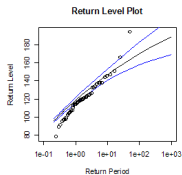
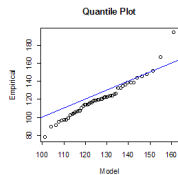
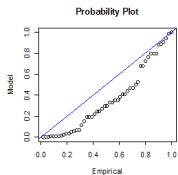
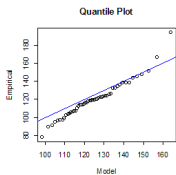
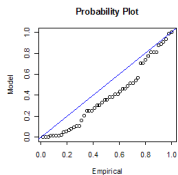


# Estadísticos de Orden

Si aplicamos este método para  $r = 1, 5$  y  $10$  obtenemos los siguientes resultados:

| $r$ | $\ell$  | $\hat{\mu}$ | $\hat{\sigma}$ | $\hat{\xi}$    |
|-----|---------|-------------|----------------|----------------|
| 1   | -222.7  | 111.1 (2.6) | 17.2 (1.8)     | -0.077 (0.074) |
| 5   | -732.0  | 118.6 (1.6) | 13.7 (0.8)     | -0.088 (0.033) |
| 10  | -1149.3 | 120.4 (1.3) | 12.7 (0.5)     | -0.115 (0.019) |

# Estadísticos de Orden



# Estadísticos de Orden

